(1)

REDUCTION OF CONSTRAINED MAXIMA

TO SADDLE-POINT PROBLEMS

Kenneth J. Arrow
Leonid Hurwicz

P-706

7 July 1955

The RAND Corporation
1700 MAIN ST. • SANTA MONICA • CALIFORNIA

REDUCTION OF CONSTRAINED MAXIMA TO SADDLE-POINT PROBLEMS

Kenneth J. Arrow and Leonid Hurwicz[1]

I. Introduction

I.1. The usual applications of the method of Lagrangian multipliers, used in locating constrained extrema (say maxima), involve the setting up of the Lagrangian expression

$$\phi(x,y) = f(x) + y'g(x) \tag{1}$$

where $f(x)$ is being (say) maximized with respect to the (vector) variable x, subject to the constraint (vector) $g(x) = 0$ and y is the Lagrange multiplier (vector).[2]

The essential step of the customary procedure is the solution for x, as well as y, of the pair of (vector) equations,[3]

$$\phi_x(x,y) = 0 \tag{2.1}$$

$$g(x) = 0. \tag{2.2}$$

Let $(\bar{x},\bar{y})$ be the solution of equations (2), while $\hat{x}$ maximizes $f(x)$ subject to $g(x) = 0$. Then, under suitable restrictions,

$$\bar{x} = \hat{x}. \tag{3}$$

---

1. The authors wish to thank The RAND Corporation, under whose auspices most of this work was done, and the Office of Naval Research for additional support and assistance.

2. $x = \left\{x_1,\ldots,x_N\right\}$, $y = \left\{y_1,\ldots,y_M\right\}$. $\left\{\quad\right\}$ indicates a column vector. The prime indicates transposition; thus y' is a row vector. $g(x)$ maps the points of the N-dimensional x-space into an M-dimensional space; we have
$g(x) = \left\{g_1(x),\ldots,g_M(x)\right\}$.

3. $\phi_x(x,y) = \left\{\frac{\partial\phi(x,y)}{\partial x_1},\ldots,\frac{\partial\phi(x,y)}{\partial x_N}\right\}$.

I.2. In [1] Kuhn and Tucker treat the related problem of maximizing $f(x)$ subject to the constraints[4,5] $g(x) \geqq 0$, $x \geqq 0$. This problem includes, by a suitable reformulation, the case of maximizing $\underline{f}$ subject to both equalities and inequalities, since (g_m being real-valued) the equality $g_m(x) = 0$ is equivalent to the pair of inequalities $g_m(x) \geqq 0$, $-g_m(x) \geqq 0$. The case where a component x_n of $\underline{x}$ is unrestricted as to sign can also be handled within this framework through a replacement of x_n by the difference $x_{n1}-x_{n2}$ where the two new variables x_{n1}, x_{n2} are required to be non-negative. (In later sections of this paper we shall treat directly the class of situations where $f(x)$ is to be maximized subject to $g^{(1)}(x) \geqq 0$, $g^{(2)}(x) = 0$, $x^{[1]} \geqq 0$, $x^{[2]}$ not restricted as to sign, $x = \{x^{[1]}, x^{[2]}\}$.)

Denote by C_g the set of all $\underline{x}$ satisfying the constraints $g(x) \geqq 0$, $x \geqq 0$. The following two results are of fundamental importance in what follows.

A. ([1]. Theorem 1.) Let $\underline{g}$ satisfy the following condition (called Constraint Qualification, here abbreviated as C.Q.):[6] If $\tilde{x}$ is a boundary point of C_g, and $\underline{x}$ satisfies the relations

$$(4.0) \qquad \begin{cases} (4.0') & \tilde{g}_x^a(x-\tilde{x}) \geqq 0, \\ (4.0'') & x^b-\tilde{x}^b \geqq 0, \end{cases}$$

4. For an arbitrary K-dimensional vector $a = \{a_1, a_2, \ldots, a_K\}$ the relation $a \geqq 0$ is here defined to mean $a_k \geqq 0$ for $k = 1,2,\ldots,K$. (Another definition of such vectorial inequalities, permitting greater generality of treatment, will be used below.) $a > 0$ means $a_k > 0$ for $k = 1,2,\ldots,K$.

5. In [1] our $\underline{f}$ and $\underline{g}$ are respectively written as $\underline{g}$ and $\underline{F}$. The symbol in [1] for the Lagrange multiplier (our $\underline{y}$) is $\underline{u}$.

6. [1, p. 483]. This restriction "is designed to rule out singularities on the boundary of the constraint set, such as an outward-pointing 'cusp'." It should be noted, however, that (because of (4.0')) C.Q. is a property of $\underline{g}$, not merely of C_g. Thus $g(x) \equiv -(x-1)^3$, $\underline{x}$ one-dimensional, lacks C.Q., while $g(x) \equiv -(x-1)$, with the same C_g, does have it.

where "$\sim$" over a symbol denotes its evaluation at $x = \tilde{x}$, $g = \{g^a, g^\beta\}$, $\tilde{g}^a = 0$, $\tilde{g}^\beta > 0$, $x = \{x^a, x^b\}$, $\tilde{x}^a > 0$, $\tilde{x}^b = 0$, then there exists a differentiable vector-valued function Ψ of the real variable θ whose domain is the closed interval $[0,1]$ and the range is in C_g, i.e., $x = \Psi(\theta)$, such that $\Psi(0) = x$ and $\Psi'(0) = \lambda(x-\tilde{x})$ for some positive scalar λ.

Under this condition, if all derivatives used below exist and if $\bar{x}$ maximizes $f(x)$ for $x \epsilon C_g$, there exists y satisfying the conditions

$$(4.1) \qquad \bar{x} \geqq 0, \ \bar{\phi}_x \leqq 0, \ \bar{x}'\bar{\phi}_x = 0,$$

$$(4.2) \qquad \bar{y} \geqq 0, \ \bar{\phi}_y \geqq 0, \ \bar{y}'\bar{\phi}_y = 0,$$

where $\bar{\phi}_x = \left\{\frac{\partial\phi(x,y)}{\partial x_1}, \dots, \frac{\partial\phi(x,y)}{\partial x_N}\right\}$ and $\phi_y = \left\{\frac{\partial\phi(x,y)}{\partial y_1}, \dots, \frac{\partial\phi(x,y)}{\partial y_M}\right\}$ are partial (vector) derivatives of the Lagrangian expression (1) evaluated at $(\bar{x},\bar{y})$.

B. ([1]. Theorem 3.) If the hypotheses specified in (A) hold and, in addition, the functions $f(x)$, $g_m(x)$, $m = 1,\dots,M$ are <u>concave</u>,[7] there exists a pair $(\bar{x},\bar{y})$, satisfying conditions (4), such that (x,y) is a <u>non-negative saddle-point</u> (NNSP) of $\phi(x,y)$, i.e.,

$$(5) \qquad \phi(x,\bar{y}) \leqq \phi(\bar{x},\bar{y}) \leqq \phi(\bar{x},y) \text{ for all } x \geqq 0, \ y \geqq 0;$$

furthermore, any NNSP $(\bar{x},\bar{y})$ of $\phi(x,y)$ has the property that $\bar{x}$ maximizes $f(x)$ in C_g.[8]

---

7. A function $f(x)$ is said to be concave if

$$(1-\theta)f(x^\circ) + \theta f(x) \leqq f\left[(1-\theta)x^\circ + \theta x\right]$$

for all $0 \leqq \theta \leqq 1$ and all x° and $\underline{x}$ in the region where $f(x)$ is defined. See [1, pp. 10-11].

8. According to Lemma 1, [1], conditions (4) are implied by (5) regardless of the nature of $\phi(x,y)$, i.e., even if $\phi(x,y)$ is not given by (1).

I.3. Game interpretation. The result contained in (B) above can be interpreted in the language of the theory of games. It states that it is possible to set up a two-person zero-sum game g_ϕ where the maximizing player controls $x \geqq 0$ while the minimizing player controls $y \geqq 0$ and the payoff to the maximizing player is given by $\phi(x,y) = f(x) + y'g(x)$. Mixed strategies are regarded as excluded. (B) implies that the solution $\hat{x}$ of the constrained maximization problem can be obtained as an optimal strategy[9] $\bar{x}$ of the maximizing player in g_ϕ; furthermore, any optimal strategy $\bar{x}$ of the maximizing player in g_ϕ has the property $\bar{x} = \hat{x}$. It will be noted that the corresponding optimal strategy $\bar{y}$ of the minimizing player yields the Lagrangian multiplier for the constrained maximization problem.

I.4. Alternative game interpretation. Let Γ be a given (vector) constant and define $h(x)$ by

$$h(x) = \Gamma - g(x). \tag{7}$$

Then the Lagrangian expression (1) becomes

$$\phi(x,y) = f(x) + y'\left[\Gamma - h(x)\right]. \tag{8}$$

It is now possible to set up a two-person game g_ϕ^* (which is not necessarily zero-sum) where the payoff function of the x-player is

$$\pi^x(x,y) = f(x) - y'h(x) \tag{9}$$

---

9. A pair of strategies $(\bar{x},\bar{y})$ is said to be optimal in a zero-sum two-person game if $\bar{x} \in X$, $\bar{y} \in Y$ where X,Y, are the respective strategy domains of the two players and the relation

$$P(x,\bar{y}) \leqq P(\bar{x},\bar{y}) \leqq P(\bar{x},y) \qquad \text{for all } x \in X,\ y \in Y$$

holds, where $P(x,y)$ is the payoff function of the maximizing x-player.

while the payoff function of the y-player is

$$(10) \qquad \pi^{y}(x,y) = -\, y'\left[\zeta - h(x)\right].$$

With reference to g_{ϕ}^{*}, the result of (B) above implies that the solution $\hat{x}$ of the constrained maximization problem can be obtained as an optimal strategy[10] $\bar{x}$ of the x-player in g_{ϕ}^{*}.

As before, any optimal $\bar{x}$ strategy of the x-player has the property $\bar{x} = \hat{x}$. Also, the optimal strategy of the y-player, $\bar{y}$, gives the Lagrange multiplier of the Lagrange problem.

I.5. Economic interpretation. The economic interpretation of (B), as suggested by the authors of [1], can best be stated in relation to the game g_{ϕ}^{*}. Let x (by definition non-negative) be the **activity level** vector (cf. Koopmans, [2]). Interpret ζ as the available (vector) amount of primary commodities while $z = h(x)$ is the (vector) amount actually used. Denote by $v = \{v_1, \ldots, v_p\}$ the desired commodity output (vector). We have $v = k(x) = \{k_1(x), \ldots, k_P(x)\}$. Finally, assume (for the sake of simplicity) that the utility function is linear in the components of $\underline{v}$, say $u(v) = \sum_{p=1}^{P} a_p v_p = a'v$

---

10. Using the concept proposed by Nash, optimality in an N-person game is defined as follows. Write x_n for the strategy of the n-th player, $x_n \in X_n$ (i.e., $x \in X$); let $x = \{x_1, \ldots, x_N\}$ and $x_{)n(} = \{x_1, \ldots, x_{n-1}, x_{n+1}, \ldots, x_N\}$. Denote the payoff function of the n-th player by $P_n(x_n, x_{)n(})$. Then $\bar{x} \in X$ is said to be an equilibrium point of the game and the components of $\bar{x}$ optimal if

$$P_n(\bar{x}_n, \bar{x}_{)n(}) = \max_{x_n \in X_n} P_n(x_n, \bar{x}_{)n(}) \text{ for } n = 1, \ldots, N.$$

It will be noted that the saddle-point definition of optimality for a two-person zero-sum game (see preceding footnote) is a special case of the present (Nash) definition.

with $\alpha = \{\alpha_1, \ldots, \alpha_p\}$. Define $f(x) = u(v)$ where $\underline{v}$ is a function of $\underline{x}$, i.e.,

$$f(x) = \sum_{p=1}^{P} \alpha_p k_p(x) = \alpha' k(x). \tag{11}$$

Then the problem of maximizing utility with regard to the activity level, subject to availability of the required primary commodities, requires the maximization of $f(x)$ subject to $x \geq 0$, $g(x) \geq 0$ where $g(x)$ is obtained from $\int$ and $h(x)$ according to (7).

The "allocation game" $g^*_{\phi A}$ (cf. Koopmans, [2] is played by three persons: the helmsman, the manager, and the custodian. The helmsman sets (once and for all) the desired commodity price vector α; he also sells to the custodian the available amount $\int$ of the primary commodities. The custodian sets the price vector of the primary commodities $\underline{y}$. He purchases $\int$ from the helmsman at this price and sells an amount $\underline{z}$ to the manager at the same price. The manager purchases $\underline{z}$ from the custodian, determines the level of activity $\underline{x}$ which will just use up $\underline{z}$, produces the amount $\underline{v}$ of desired commodities and sells this to the helmsman at the present price α.

While three persons appear in the description of $g^*_{\phi A}$, the role of the helmsman is purely auxiliary (equivalent to formulating certain rules of the game). Hence we shall regard this game as a two-person game, the two players being the manager (controlling $\underline{x}$) and the custodian (controlling $\underline{y}$). The payoff functions for each are given by the net profit on the transactions undertaken. The revenue and cost of the two players may be presented in a tabular form:

	revenue	cost
Manager	$f(x) = \alpha' k(x)$	$y' h(x)$
Custodian	$y' h(x)$	$y' \int$

Clearly $g^*_{\emptyset A}$ is a special case of $g^*_{\emptyset}$ for the $f(x)$ specified by (11). Hence a solution of $g^*_{\emptyset A}$ will have the properties described in I.4 implied by (B). In particular, the activity level maximizing the manager's profits will also maximize the utility function subject to the availability of primary commodities.

I.6. Decentralization. A case of particular interest arises when additivity holds, viz.

$$(12) \qquad k_p(x) = \sum_{n=1}^{N} k_p^n(x_n) \qquad p = 1,2,\dots,P$$

$$h_m(x) = \sum_{n=1}^{N} h_m^n(x_n) \qquad m = 1,2,\dots,M\ .$$

("Linear programming" is a special case of (12) with

$$(13) \qquad k_p^n(x_n) = \delta_p^n x_n \qquad p = 1,2,\dots,P$$

$$h_m^n(x_n) = \gamma_m^n x_n \qquad m = 1,2,\dots,M$$

where δ_p^n and γ_m^n are constants.)

When (12) holds a decentralized allocation game $g^{*D}_{\emptyset A}$ may be obtained from $g^*_{\emptyset A}$ as follows. Each component of $\underline{y}$ is now controlled by a separate custodian who deals in only one primary commodity and whose payoff function is given by

$$(14) \qquad \pi^{y_m}(x,y) = -\, y_m\Big(\rho_m - \sum_{n=1}^{N} h_m^n(x_n)\Big) \qquad m = 1,2,\dots,M.$$

Similarly each component of $\underline{x}$ is controlled by a separate manager who is in charge of one plant only and whose payoff function is given by

$$(15) \qquad \pi^{x_n}(x,y) = \sum_{p=1}^{P} a_p\, k_p^n(x_n) - \sum_{m=1}^{M} y_m h_m^n(x_n) \qquad n = 1,2,\dots,N.$$

As before, these payoff functions may be interpreted as net profits on the transactions undertaken by the custodians and managers.

It is again true that the optimal (for definition, see footnote in I.4 above) strategies (x,y) of the $(N + M)$ - person non-zero-sum game $\mathcal{G}^{*D}_{A}$ are such that $\bar{x} = \hat{x}$ and $\bar{y}$ is the Lagrangian multiplier.

II. A Modified Lagrangian Approach.

II.1. Because of the interesting game theoretical and economic implications of the Theorem in (B), I.2 above (some of which the authors will study elsewhere), the question arises as to the possibility of similar results when some of the conditions of the theorem are relaxed.

It turns out that results of such nature can be obtained, though not without some sacrifices. The relaxation is primarily with regard to the convexity assumptions which fail to hold in some important economic applications (the case of "increasing returns"). The main sacrifices are these: the Lagrangian expression is modified and the results are proved only locally.

The results are presented below in the form of three theorems. Some remarks concerning the interpretation of results are contained in III below.

Theorem 1 is auxiliary in nature; Theorems 2 and 3 together imply the existence of a local non-negative saddle-point for the modified Lagrangian expression. Theorem 3 shows this saddle-point to be of the type leading to convergence in gradient procedures described by the authors in [2].

The notation differs in some detail from that introduced in I above. To facilitate reading some notational principles, principles are stated in II.2.1 and the main symbols used are listed in 2.2 and 2.3.

II.2.1. Some Principles of Notation. A K-dimensional column vector $\{a_1, a_2, \ldots, a_K\}$ is denoted by $\underline{a}$; dim a denotes the number of components in $\underline{a}$. If A is a matrix, A' is its transpose. Hence, in particular, a' is a row vector and $a'b = \sum_{k=1}^{K} a_k b_k$ is the inner product of the vectors $\underline{a}$ and $\underline{b}$; a.b is an alternative, and sometimes more convenient, notation for a'b.

$[a_1, a_2, \ldots, a_K]$ is the finite (unordered) set whose elements are $a_1, a_2, \ldots, a_K$. $A \sim B$ is the set of all elements in A but not in B (the set-theoretic difference).

$\{x | P_x\}$ denotes the set of all $\underline{x}$ possessing the property P_x. If

$$c(a) = \{c_1(a), c_2(a), \ldots, c_P(a)\},$$

$$a = \{a_1, a_2, \ldots, a_K\},$$

then

$$c_a \equiv c_a(a) = \left\| \frac{\partial c_p}{\partial a_k} \right\|, \qquad p = 1,2,\ldots,P; \quad k = 1,2,\ldots,K .$$

$\bar{c}$, $\bar{c}_a$ denote, respectively, c(a) and $c_a(a) \equiv c_a$ evaluated at $a = \bar{a}$.

If $\Psi(a,b)$ is a real-valued (scalar) function of the vectors

$$a = \{a_1, a_2, \ldots, a_K\} \qquad b = \{b_1, b_2, \ldots, b_R\},$$

then

$$\Psi_{ab} = \left\| \frac{\partial^2 \Psi}{\partial a_k \, \partial b_r} \right\| , \quad k = 1,2,\ldots,K; \quad r = 1,2,\ldots,R .$$

$\bar{\Psi}_{ab}$ denotes Ψ_{ab} evaluated at $(\bar{a}, \bar{b})$.

$S_\rho(x^o) = \{x | d(x,x^o) \leq \rho\}$ where $d(x',x'')$ denotes the Euclidean distance between x' and x''.

II.2.2. Some Symbols Used.

(N.1.1) $x = \{x_1, x_2, \ldots, x_N\}$.

X is the Euclidean N-space of the x's.

$\mathcal{N} = 1, 2, \ldots, N$.

$\mathcal{N}'$ is a fixed (possibly empty, not necessarily proper) subset of $\mathcal{N}$. (As will be seen in (N.1.4) below, the elements of $\mathcal{N}'$ are the indices of the components of $x^{(1)}$ as defined in the first paragraph of I.2.)

(N.1.2) $z = \{z_1, z_2, \ldots, z_M\}$.

Z is the Euclidean M-space of the z's.

$\mathcal{M} = [1, 2, \ldots M]$.

$\mathcal{M}'$ is a fixed (possibly empty, not necessarily proper) subset of $\mathcal{M}$. (As will be seen from (N.1.4), (N.2), (N.3) below, the elements of $\mathcal{M}'$ are the indices of the components of $g^{(1)}$ as defined in the first paragraph of I.2; the elements of $\mathcal{M} \sim \mathcal{M}'$ are the indices of $g^{(2)}$ (cf. ibid.); g will be defined as $\{g^{(1)}, g^{(2)}\}$.)

(N.1.3) $y = \{y_1, y_2, \ldots, y_M\}$.

Y is the Euclidean M-space of the y's.

(Y is the space of the real-valued linear functions on Z. Even in the Euclidean case it is convenient to distinguish between the two, since our definitions of non-negativity in the two spaces differ.)

(N.1.4) $x \geqq 0$ means $\begin{cases} x_n \geqq 0 & \text{for } n \in \mathcal{N}'. \\ x_n \text{ unrestricted as to sign for } n \notin \mathcal{N}'. \end{cases}$

X^+ is the set of all $x \geqq 0$.

$z \geqq 0$ means $\begin{cases} z_m \geqq 0 & \text{for } m \in \mathcal{M}'. \\ z_m = 0 & \text{for } m \notin \mathcal{M}'. \end{cases}$

$y \geqq 0$ means $\begin{cases} y_m \geqq 0 & \text{for } m \in \mathcal{M}'. \\ y_m \text{ unrestricted as to sign for } m \notin \mathcal{M}'. \end{cases}$

For any vector $a = \{a_1, a_2, \ldots, a_K\}$,

$a = 0$ means $a_1 = 0, a_2 = 0, \ldots, a_K = 0$;

$a > 0$ means $a_1 > 0, a_2 > 0, \ldots, a_K > 0$;

$a < 0$ means $-a > 0$.

(N.2.1) $'g$ is a function on X^* to Z. Hence

$'g(x) = \{'g_1(x), 'g_2(x), \ldots, 'g_M(x)\}$ where the $'g_m$, $m \in \mathcal{M}$, are real-valued functions.

(N.2.2) We shall find it convenient to work with some of the $'g_m$, $m \in \mathcal{M}'$, replaced by their negatives.

More precisely, we write

$$g_m = \begin{cases} 'g_m & \text{if } m \in \mathcal{M} \sim \mathcal{M}^- \\ -'g_m & \text{if } m \in \mathcal{M}^- \end{cases}$$

where $\mathcal{M}^- \subset \mathcal{M} \sim \mathcal{M}'$ is defined in II.3.4C below.

$g = \{g_1, g_2, \ldots, g_M\}$.

<u>Note</u>. Since $\mathcal{M}^- \subset \mathcal{M} \sim \mathcal{M}'$, it is seen that the conditions

$$'g(x) \geqq 0, \quad g(x) \geqq 0$$

are equivalent. For practical purposes, one could consider the problem as given directly in terms of g, rather than $'g$. We start with $'g$, however, in order to avoid the impression of a loss of generality in connection with the assumptions of II.3.4.C.

(N.3) $C_g = \{x \mid g(x) \geq 0,\ x \geq 0\} = \{x \mid g(x) \geq 0,\ x \geq 0\}$

(the "constraint set").

(N.4) f is a real-valued function on X^* (the "maximand").

(N.5) $O_{fg} = \{x' \mid x' \in C_g$ and $f(x) \leq f(x')$ for all $x \in C_g\}$

(the "optimal set").

(N.6) $x = \{x^{(1)}, x^{(2)}\}$ where

$\mathcal{N}^{(i)}$ = the set of indices of the components of $x^{(i)}$, $i = 1,2$

$n \in \mathcal{N}^{(1)}$ if $n \notin \mathcal{N}'$ or $n \in \mathcal{N}'$ and $\bar{x}_n > 0$

$n \in \mathcal{N}^{(2)}$ if $n \in \mathcal{N}'$ and $\bar{x}_n = 0$

for a given $\bar{x} \in O_{fg}$ and either component may be empty.

Note 1. When a vector a is partitioned into two subvectors, say

$$a = \{a^*, a^{**}\}$$

and we say that a^* (or: a^{**}) is empty, this means that $a = a^{**}$ (or: $a = a^*$).

Note 2. The above partitioning of the vector x obviously depends on the point $\bar{x}$ in O_{fg} chosen. The same is true of the partitioning in (N.7) below and of various subsequent partitionings of x and g. It is understood that all these partitionings refer to the same choice of $\bar{x}$, and that $\bar{x}$, once chosen, remains fixed.

(N.7) $g = \{g^{(1)}, g^{(2)}\}$

where

$$g^{(1)}(\bar{x}) = 0, \quad g^{(2)}(\bar{x}) > 0$$

and either component may be empty.

(N.8) $h(x) = 1 - g(x)$

where 1 denotes the M-dimensional vector with 1's as components; $h^{(i)} = 1 - g^{(i)}$, $i = 1,2$.

(N.9) $\quad {}_{\eta_m}p_m(x) = 1 - \left[h_m(x)\right]^{1+\eta_m} \qquad m \in \mathcal{M}.$

(N.10) $\quad \eta = \{\eta_1, \eta_2, \ldots, \eta_M\}.$

(N.11) $\quad {}_{\eta}p(x) = \{\eta_1 p_1(x), \eta_2 p_2(x), \ldots, \eta_M p(x)\}.$

(N.12) $\quad {}_{\eta}\phi(x,y) = f(x) + y'({}_{\eta}p(x))$ (the "modified Lagrangian expression").

II.3.1.1. A Reformulation of Kuhn-Tucker Theorem 1.

This slight generalization of Theorem 1 in [1, p. 484] is needed here because of the meaning of inequalities given in (N.1.4). [The possibility of this type of generalization is indicated in [1, Sec. 8, pp. 491-2].][11]

We shall say that g satisfies the Constraint Qualification (C.Q.) at x, if the requirements of the definition in I.2.A are satisfied with the inequalities (4.0) in I.2.A interpreted in the sense of (N.1.4). $\phi(x,y)$ is given by (1) in I.1. (It is immaterial whether g or $'g$ is used.)

Theorem.

If f and g are differentiable, $\bar{x} \in O_{fg}$ and g satisfies C.Q. at $\bar{x}$, then there exists a $\bar{y} \in Y$ such that

$$\bar{y} \geqq 0; \quad \bar{\phi}_y \cdot \bar{y} = 0; \quad \bar{\phi}_y \cdot y \geqq 0 \qquad \text{for all} \quad y \geqq 0;$$

$$\bar{x} \geqq 0; \quad \bar{\phi}_x \cdot \bar{x} = 0; \quad \bar{\phi}_x \cdot x \leqq 0 \qquad \text{for all} \quad x \geqq 0\,.$$

[Note that, by virtue of the definitions in II.2.2, this means that $\bar{\phi}_{y_m} \geqq 0$ if $m \in \mathcal{M}'$, $\bar{\phi}_{y_m} = 0$ if $m \notin \mathcal{M}'$, $\bar{\phi}_{x_n} \leqq 0$ if $n \in \mathcal{N}'$, $\bar{\phi}_{x_n} = 0$ if $n \notin \mathcal{N}'$. The other inequalities of the theorem are also to be interpreted in the sense of (N.1.4).]

---

11. See also Hurwicz, [9, pp. VIII - 2-6].

II.3.1.2. Theorem 1.

Definition.[12] An M-dimensional vector $\eta = \{\eta_1, \eta_2, \ldots, \eta_M\}$ is said to be acceptable if, for each $m \in \mathcal{M}$, (1) $\eta_m \geqq 0$, and (2) η_m is an even integer if $h_m(\bar{x}) < -1$.

Theorem.

If, for some $\rho > 0$, $x \in S_\rho(\bar{x})$, $\bar{x} \in O_{fg}$, f and g are differentiable, and g satisfies $C_{\cdot q}$ at $\bar{x}$, then, for any acceptable η, there exists a vector $\bar{y} = \bar{y}(\eta)$ such that[13]

(I.1') $\quad \eta\bar{\phi}_x . x \leqq 0 \qquad$ for all $\quad x \geqq 0$;

(I.1") $\quad \eta\bar{\phi}_x . \bar{x} = 0$;

(I.1"') $\quad \bar{x} \geqq 0$;

(I.2') $\quad \eta\bar{\phi}_y . y \geqq 0 \qquad$ for all $\quad y \geqq 0$;

(I.2") $\quad \eta\bar{\phi}_y . \bar{y} = 0$;

(I.2"') $\quad \bar{y} \geqq 0$.

(The relations (I) are necessary conditions for a nonnegative (in the sense of (M.1.4)) saddle-point of $\eta\phi(x,y)$ at $(\bar{x},\bar{y})$.)

In particular, the relations (I) are satisfied if one selects $\bar{y} = \bar{y}(\eta)$ such that[14]

(I.3') $\quad (1+\eta_m)\bar{y}_m(\eta) = \bar{y}_m(0) \qquad$ for all $\quad m \in \mathcal{M}$.

---

12. In many applied problems, $h_m(x) \geqq 0$ for all m and all $x \geqq 0$. It was pointed out by Dr. Masao Fukuoka that, in the absence of such an assumption, the requirement of non-negativity of the components of η is insufficient for the proof of the theorem.

13. The bar over ϕ denotes evaluation at $x = \bar{x}$, $y = \bar{y}(\eta)$.

14. $\bar{y}_m(0) = \bar{y}$ in Kuhn-Tucker Theorem 1 (cf. II.3.1.1).

If the selection is made in accordance with (I.3'), the equality

(I.3") $$ {}_0\phi_x = {}_\eta\phi_x $$

will hold. (${}_0\phi(x,y)$ is ${}_\eta\phi(x,y)$ with $\eta = 0$; this is obviously the same as $\phi(x,y)$ in (1) of I.1.)

Proof.

For $\eta = 0$, the preceding theorem follows directly from the reformulated version of the Kuhn-Tucker Theorem 1 given in II.3.1.1. Thus there exists a vector

(1') $$ \bar{y}(0) = \{\bar{y}_1(0), \bar{y}_2(0), \ldots, \bar{y}_M(0)\} $$

with the required properties.

Consider now the case $\eta \neq 0$. We shall show that $\bar{y}(\eta)$ defined by (I.3'), i.e., explicitly, by

(1") $$ y_m(\eta) = \frac{1}{1 + \eta_m} \bar{y}_m(0), \qquad m \in \mathcal{M} $$

(where $\bar{y}_m(0)$ is that of (1')) satisfies the relations (I).

We first observe that (1") yields

(2) $$ (1 + \eta_m)\, \bar{y}_m(\eta) \left[h_m(\bar{x})\right]^{\eta_m} = \bar{y}_m(0), \qquad m \in \mathcal{M}. $$

(When $h_m(\bar{x}) = 1$, (2) follows directly from (1"). When $h_m(\bar{x}) \neq 1$, we have ${}_0\phi_{y_m} = g_m(\bar{x}) > 0$, and hence, by (I.2), $\bar{y}_m(0) = 0$; (1") then yields $\bar{y}_m(\eta) = 0$ and (2) follows.)

Since

(3) $$ {}_\eta\phi_{x_n} = f_{x_n} + \sum_{m=1}^{M} (1 + \eta_m) y_m(\eta)\, h_m(x)^{\eta_m} \frac{\partial g_m(x)}{\partial x_n}, \qquad n \in \mathcal{N}, $$

(2) implies

$$(4) \qquad {}_{\eta}\bar{\phi}_{x_n} = \bar{f}_{x_n} + \sum_{m=1}^{M} \bar{y}_m(0) \frac{\partial g_m(\bar{x})}{\partial x_n}, \qquad n \in \mathcal{N}.$$

Noting that the right member of (4) is ientical with ${}_{o}\bar{\phi}_{x_n}$, we conclude that the relations (I.1) hold for all η with non-negative components, since they are known to hold for $\eta = 0$.

(I.2') is established by the fact that the right member of

$$(5) \qquad {}_{\eta}\bar{\phi}_{y_m} = {}_{\eta}p_m(\bar{x}) = 1-\left[h_m(\bar{x})\right]^{1+\eta_m}, \qquad m \in \mathcal{M},$$

is non-negative for $m \in \mathcal{M}'$, zero for $m \notin \mathcal{M}'$ when η is acceptable (see the definition above) since, for any $m \in \mathcal{M}$, $h_m(\bar{x}) \leqq 1$, and, furthermore, $\bar{\phi}_{y_m} = 0$ if $m \notin \mathcal{M}'$, in which case $h_m(\bar{x}) = 1$.

Now suppose that, for some $m_o \in \mathcal{M}$, ${}_{\eta}\bar{\phi}_{y_{mo}} > 0$, i.e., $h_{m_o}(\bar{x}) < 1$; then, by (I) for $\eta = 0$, $\bar{y}_{m_o}(0) = 0$; hence $\bar{y}_{m_o}(\eta) = 0$, and, therefore,

$$(6) \qquad \bar{\phi}_{y_{mo}} \cdot \bar{y}_{m_o}(\eta) = 0.$$

Since (6) clearly holds in the alternative case ${}_{\eta}\bar{\phi}_{m_o} = 0$, (I.2") follows.

Finally, (I.2'") holds because $\bar{y}_m(\eta)$ has the same sign as $\bar{y}_m(0)$ and the latter, by (I.2'") for $\eta = 0$, is non-negative if $m \in \mathcal{M}'$.

II.3.2. <u>Theorem</u> 2.

Let, for some $\varsigma > 0$, $x \in S_{\varsigma}(\bar{x})$, $\bar{x} \in O_{fg}$, such that (I) is satisfied. Then

$$(II) \qquad {}_{\eta}\phi(\overline{x,y}) \leqq {}_{\eta}\phi(\bar{x},y) \qquad \text{for all} \quad y \geqq 0.$$

For we have

$${}_{\eta}\phi(\bar{x},y) - {}_{\eta}\phi(\bar{x},\bar{y}) = (y-\bar{y}) \cdot {}_{\eta}\bar{p} = y \cdot {}_{\eta}\bar{p} \geqq 0 \qquad \text{for } y \geqq 0$$

where, since

$${}_\eta\bar{\phi}_y = {}_\eta\bar{p}\ ,$$

the second equality follows from (I.2") and the inequality from (I.2').

II.3.3. Notation.

(N.13) $$x^{(2)} = \left\{ x^{(21)},\ x^{(22)} \right\}$$

where

$${}_o\bar{\phi}_{x^{(21)}} = 0, \qquad {}_o\bar{\phi}_{x^{(22)}} < 0\ ,$$

and either component may be empty.

(N.14) $$x = \left\{ x^{I},\ x^{II} \right\}$$

where

(N.14*) $$x^{I} = \left\{ x^{(1)},\ x^{(21)} \right\}$$

$$x^{II} = x^{(22)}.$$

(Either x^{I} or x^{II} may be empty.)

It should be noted that, by (I) and (N.13),

(N.14**) $${}_o\bar{\phi}_{x^{I}} = 0\ ,$$

$${}_o\bar{\phi}_{x^{II}} < 0\ .$$

II.3.4. Definition of a Regular Constrained Maximum.

A. In Theroem 3 below we use the concept of a regular constrained maximum. The definition of such a maximum is given in II.3.4E. To state it, we must first formulate three regularity conditions: $\overset{1}{\bigtriangledown}$, $\overset{2}{\bigtriangledown}$, $\overset{3}{\bigtriangledown}$ (eqs. (B.8) or (C.10), (C.8), (D.19)).

B. The First Regularity Condition.

Let $\bar{x}$ be a value maximizing the function $f(x)$ subject to $g(x) \geqq 0$, $x \geqq 0$, and hence also subject to

(B.1) $$g(x) \geqq 0$$
$$x \geqq 0$$

where the inequalities are to be interpreted in the sense of (N.1.4).

From (N.6) and (N.7) it is clear that, for sufficiently small variations of $\underline{x}$, the constraints

(B.2) $$g^{[2]}(x) \geqq 0$$
$$x^{(2)} \geqq 0.$$

which are a part of (B.1), can be disregarded. Hence, at $\bar{x}$, $f(x)$ possesses a local maximum subject to

(B.3) $$g^{[1]}(x) \geqq 0,$$
$$x^{(2)} \geqq 0.$$

Let g^t be a subvector of $g^{[1]}$ such that $C_g = C_{\{g^t, g^{[2]}\}}$ and write

(B.4) $$g^{[1]} = \{g^t, g^{tt}\}.$$

The components of g^{tt}, unless they produce an inconsistency,[15] can be disregarded in the process of maximization. (I.e., $O_{f,g} = O_{f,\{g^t,g^{[2]}\}}$.) If the Lagrangian multiplier vector $\bar{y}^{[1]}$ (corresponding to the constraints $g^{[1]}(x) \geqq 0$) is partitioned according to

(B.5) $$\bar{y}^{[1]} = \{\bar{y}^t, \bar{y}^{tt}\},$$

it is always possible to put

(B.6) $$\bar{y}^{tt} = 0 ,$$

---

15. Which cannot happen since $\bar{x}$ exists.

and this will be done in what follows.

Assuming that the constraints (B.3) are consistent, we may replace them by

$$\text{(B.7)} \qquad \begin{aligned} & g^t(x) = 0 \\ & x^{(2)} \geqq 0 \,. \end{aligned}$$

The <u>first regularity condition</u> $\bigtriangledown\!\!\!\!\!\!\;1$ is

$$\text{(B.8)} \qquad \operatorname{rank}\,(\bar{g}^t{}_x(1)) = \dim g^t = M^t \,,$$

say.

<u>Note 1.</u> $\bigtriangledown\!\!\!\!\!\!\;1$ corresponds to the requirement of non-degeneracy in linear programming (see Dantzig, [4, p. 340]).

<u>Note 2.</u> $\bigtriangledown\!\!\!\!\!\!\;1$ implies C.Q. (see Appendix 1).

C. <u>The Second Regularity Condition.</u>

Since, by (N.4), (N.6), (N.14*), (B.4),

$$\text{(C.1)} \qquad \begin{cases} g^t(\bar{x}) = 0 \\ \bar{x}^{II} = 0 \,, \end{cases}$$

it follows that, as a function of x^I,

$$f(x^I, \bar{x}^{II}) \equiv f(x^I, 0)$$

has at $\bar{x}^I$ a local maximum subject to the constraints

$$\text{(C.2)} \qquad \begin{cases} g^t(x^I, \bar{x}^{II}) \equiv g^t(x^I, 0) = 0 \\ \qquad x^{(21)} \geqq 0 \,. \end{cases}$$

The corresponding Lagrangian expression becomes

$$\text{(C.3)} \qquad \phi^I(x^I, y^t) = f(x^I, 0) + y^t.g^t(x^I, 0) \,.$$

Using the reformulation of Kuhn-Tucker Theorem 1, given in II.3.1.1, we may assert the existence of a $\bar{y}^t$ such that

(C.4)$^{-1}$ $\quad \bar{x}^I \geqq 0; \ {}_o\bar{\phi}^I_{x^I} = 0$;

-2 $\quad \bar{y}^t \geqq 0; \ {}_o\bar{\phi}^I_{y^t} = 0$.

It might happen that some components of $\bar{y}^t$ vanish. Write

(C.6') $\quad y^t = \{y^*, y^o\}$

where

(C.6") $\quad$ every component of $\bar{y}^*$ is different from zero

and

(C.6''') $\quad \bar{y}^o = 0$.

Let g^t be correspondingly partitioned as

(C.6^{IV}) $\quad g^t = \{g^*, g^o\}$.

Now suppose that ${}_o\phi^I(x^I, y^t)$ has a non-negative saddle-point[16] at $(\bar{x}^I, \bar{y}^t)$. One can then easily verify that

(C.7) $\quad {}_o\phi^I(x^I, y^*) \equiv f(x^I, 0) + y^* \cdot g^*(x^I, 0)$

has a non-negative saddle-point at $(\bar{x}^I, \bar{y}^*)$.

But then $\bar{x}^I$ maximizes $f(x^I, 0)$ subject to $g^*(x^I, 0) \geqq 0$ and $x^{(21)} \geqq 0$. Hence in this case the components of g^o could have been disregarded in the original maximization problem $(0_{f,g} = 0_{f,\{g^*,g^{(2)}\}})$.

However, complications might arise if ${}_o\phi^I(x^I, y^t)$ did not have a

---

16. By theorem 3 in Kuhn-Tucker, a sufficient condition for this is that $\underline{f}$ and $\underline{g}$ be both concave.

non-negative saddle-point at $(\bar{x}^I, \bar{y}^t)$. To take care of this case, one might require that

(C.8') g^o is empty unless ${}_o\phi^I(x^I, y^t)$ has a local non-negative saddle-point at $(\bar{x}^I, \bar{y}^t)$.

However, to simplify matters we shall impose the seemingly[17] stronger condition

(C.8$_1'$) g^o is empty.

It follows that

(C.8$_2'$) $M^* = \dim g^* = \dim g^t = M^t$.

Let $\mathcal{M}^*$ ($= \mathcal{M}^t$ by (C.8$_1'$)) denote the set of indices of g^*. Clearly, for $m \in \mathcal{M}^* \cap (\mathcal{M} \sim \mathcal{M}')$, we may have $\bar{y}_m < 0$.

Now suppose the preceding reasoning had been carried out in terms of $'g$ instead of g. Nothing would be changed, except, possibly, the signs of some components of the Lagrangian multiplier, to be denoted by $'\bar{y}$.

I.e., we would have $'\bar{y}_m > 0$ for $m \in \mathcal{M}^* \cap \mathcal{M}'$ and $'\bar{y}_m > 0$ or $'\bar{y}_m < 0$ for $m \in \mathcal{M}^* \cap (\mathcal{M} \sim \mathcal{M}')$. Let $\mathcal{M}^-$ be defined by the relation

(C.9') $\left\{ \begin{array}{l} m \in \mathcal{M}^- \\ \text{if and only if} \\ m \in \mathcal{M}^* \cap (\mathcal{M} \sim \mathcal{M}') \\ \text{and} \\ '\bar{y}_m < 0 \; . \end{array} \right.$

Then, it is clear from (M.2.2) that we may put

(C.9") $\left\{ \begin{array}{lll} \bar{y}_m = {}'\bar{y}_m & \text{for} & m \in \mathcal{M} \sim \mathcal{M}^- \\ \bar{y}_m = -{}'\bar{y}_m & \text{for} & m \in \mathcal{M}^- \; , \end{array} \right.$

17. Cf. II.3.6.0.

so that

(C.9''') $\qquad \bar{y}_m > 0 \quad \text{for all} \quad m \in \mathcal{M}^*$.

Hence, <u>without loss of generality</u> (as compared with ($C.8_1'$)) ($C.8_1'$) may be restated as the <u>second regularity condition</u>.

$\boxed{2}$ (C.8'') $\qquad \begin{cases} g^0 \text{ is empty} \\ \text{and} \\ \bar{y}_m > 0 \quad \text{for all} \quad M \in \mathcal{M}^* . \end{cases}$

The first regularity condition then implies

(C.10') $\qquad \text{rank}\,(\bar{g}^*_{x(1)}) = M^*$

where

(C.10'') $\qquad M^* = \dim g^*$.

D. <u>The third regularity condition</u>.

D:1. When the first two regularity conditions are satisfied, second derivatives are continuous, and x^I is non-empty, it is possible to show (see Appendix 2) that a certain quadratic form is non-positive when some of the variables are restricted in sign. The third regularity condition (eq. (19)) is a strengthening of (D.18) requiring that the [illegible] form in question be negative under [illegible] restrictions. [illegible] condition, analogous [illegible] that [illegible] (, .) [illegible] second [illegible] terms [illegible].

D:2. The third regularity condition is formulated in terms of a function q(t) of a new variable vector

(D.1) $\qquad t = \{t^*, t^{**}\}$

which is obtained by a transformation of coordinates from x^I after the latter has been partitioned so that

(D.2) $\qquad x^{I} = \left\{ x^{*}, x^{**} \right\}$,

where x^* is a subvector of $x^{(1)}$.

We shall (a) define x^* and x^{**}; (b) write down the transformation defining $\left\{t^*, t^{**}\right\}$ in terms of $\left\{x^*, x^{**}\right\}$; (c) define $q(t)$; (d) formulate the third regularity condition.

In the remainder of section D it is assumed that $\boxed{\nabla 1}$ holds; in D:3 - D:4 it is also assumed that x^I is not empty.

D:3. <u>First case</u>: $M^* = 0$.

Write

(D.3) $\qquad t = t^{**} = x^{**} = x^{I}$,

so that, by (D.1) and (D.2), x^* and t^* are empty, and define (D.4)

$$\begin{aligned} q(t) &= f(x^{I}, \bar{x}^{II}) \\ &= f(t^{**}, 0). \end{aligned}$$

The third regularity condition for this case is formulated in (D.19) below.

D:4. <u>Second case</u>: $M^* > 0$.

(a) <u>The definition of</u> x^*.

From $\nabla 1$ (eq. (C.10')) it follows that there exists a (non-empty) M^*-dimensional subvector x^* of $x^{(1)}$ such that

(D.5) $\qquad \bar{g}^*_{x^*}$ is an M^* by M^* ($M^* \geqq 1$) non-singular matrix. We then define x^{**} by (D.2) and $x^{(12)}$ by

(D.6) $\qquad x^{(1)} = \left\{ x^{*}, x^{(12)} \right\}$.

Clearly

(D.7) $\qquad x^{**} = \left\{ x^{(12)}, x^{(21)} \right\}$.

(b) <u>The Transformation from</u> x^I <u>to</u> t.

Let

(D.8) $h^* = 1 - g^*$

where 1 is the M*-dimensional vector with (scalar) 1's as components. $t = \{t^*, t^{**}\}$ is then defined by the transformation

(D.9) -1) $t^* = h^*(x^*, x^{**}, x^{-II})$

-2) $t^{**} = x^{**}$.

We also partition t^{**} by

(D.10) $t^{**} = \{t^{(12)}, t^{(21)}\}$

where

$$t^{(12)} = x^{(12)}$$

$$t^{(21)} = x^{(21)}$$

This is obviously consistent with (D.[illegible]) and (D.9-[illegible])

(c) <u>The definition of</u> r(t).

By (D.8), the Jacobian H of the transformation (D.9) is

(D.12) $$H = \begin{bmatrix} h^*_{x^*} & h^*_{x^{**}} \\ 0 & I \end{bmatrix} = - \begin{bmatrix} g^*_{x^*} & g^*_{x^{**}} \\ 0 & -I \end{bmatrix},$$

so that, by (D.5),

(D.13') $$|\bar{H}| = -\left| -\bar{g}^*_{x^*} \right| \neq 0,$$

i.e.,

(D.13") $\bar{H}$ is non-singular.

Hence, locally, (D.9) can be solved for x^I in terms of <u>t</u>; we may write this solution as

(D.14) $x^I = r(t)$

where

(D.15) $r = \{r^*, r^{**}\}$

and

$$\text{(D.16)}\qquad x^{*} = r^{*}(t)$$
$$x^{**} = r^{**}(t) = t^{**}.$$

The function $q(t)$ is now defined as $f(x)$ evaluated at $x^{II} = \bar{x}^{II}$ and with x^{I} expressed in terms of $\underline{t}$, i.e.,

$$\text{(D.17)}\qquad q(t) = f(r(t), \bar{x}^{II}) \equiv f(r^{*}(t^{*}, t^{**}), t^{**}, 0).$$

D:5. <u>The statement of the third regularity condition</u>.

We have now defined $q(t)$ for all M^{*} provided the first regularity condition ($\nabla 1$) is satisfied and x^{I} is non-empty. It is shown in Appendix 2 that[18]

unless x^{**} is empty, there exists $\varsigma > 0$ such that, for all $t^{**} \epsilon S_{\varsigma}(\bar{x}^{**})$,

$$\text{(D.18)}\qquad (t^{**} - \bar{x}^{**})' \, \bar{q}_{t^{**}t^{**}} \, (t^{**} - \bar{x}^{**}) \leqq 0$$
if $t^{(21)} \geqq 0$.

The <u>third regularity condition</u> is a strengthening of the preceding inequality. It states that

$\nabla 3$

(D.19)

(D.19') x^{**} is empty

or

(D.19") there exists $\quad > 0$ such that, for all $t^{**} \epsilon S\rho(\bar{x}^{**})$.

$(t^{**} - \bar{x}^{**})' \, \bar{q}_{t^{**}t^{**}}(t^{**} - \bar{x}^{**}) < 0$

if $t^{(21)} \geqq 0$ and $t^{**} \neq \bar{x}^{**}$.

<u>Note</u>. The situation covered by (D.19') is of importance since it permits the treatment of a large class of cases where $\underline{f}$ and $\underline{g}$ are linear.

---

18. Assuming $\nabla 1$ and $\nabla 2$ and the continuity of the second derivatives.

E. Definition.

$f(x)$ is said to have a regular maximum at $\bar{x}$ subject to $g(x) \geqq 0$, $x \geqq 0$, if the three regularity conditions $\boxed{1}$, $\boxed{2}$, $\boxed{3}$ (eqs. (B.8) and hence (C.10); eq. (C.8); and eq. (D.19)), respectively are satisfied at $\bar{x}$ and $\bar{x} \in O_{fg}$.

II.3.5. Theorem 3.

If, for some $\zeta > 0$, $x \in S_\zeta(\bar{x})$, $\bar{x}$ a regular maximum,[19] of $f(x)$ subject to $g(x) \geqq 0$ and $x \geqq 0$, f and g are differentiable (with regard to x), and furthermore, when x^I is non-empty, have continuous second order derivatives with regard to x^I, then, for all acceptable[20] sufficiently large[21] η,

$$
\text{(III.1')}\quad \left\{\begin{array}{l} x^I \text{ is empty,} \\ \text{or} \\ \text{(III.1)} \quad \left\{\begin{array}{l} (x^I - \bar{x}^I)' \quad {}_\eta\phi_{x^I x^I}(x^I - \bar{x}^I) < 0 \\ \text{if } x^{(21)} \geqq 0,\ x^I \neq \bar{x}^I, \end{array}\right. \end{array}\right.
$$

and

$$
\text{(III.2')}\quad \left\{\begin{array}{l} \text{for some } \zeta' > 0, \text{ and all } x \in S_{\zeta'}(\bar{x}) \text{ such that} \\ x \geqq 0,\ x \neq \bar{x}, \\ \text{(III.2)} \quad {}_\eta\phi(x, \bar{y}(\eta)) < {}_\eta\phi(\bar{x}, \bar{y}(\eta)) \end{array}\right.
$$

---

19. The term "regular maximum" is defined in II.3.4.E.

20. The term "acceptable" is defined at the beginning of II.3.1.2.

21. "Sufficiently large η" is defined to mean that each component η_m, $m \in L$, of η is sufficiently large.

where $\eta\phi$ and $\bar{y}(\eta)$ are defined as in Theorem 1.

Note.[22] Theorem 3 is valid for f,g linear if x^{**} is empty (regardless of whether x^* is empty), provided the first two regularity conditions hold. [If both x^* and x^{**} are empty, x^I is empty, and the theorem follows from II.3.6.A(a). If x^{**} is empty while x^* is non-empty, use II.3.6.A(a) and (b) together with II.3.6.B.2.1-3 (since g^* is non-empty and t^{**} is empty).]. Note that x^{**} is empty at the basic solutions of a linear programming problem.

II.3.6. Proof of Theorem 3.

II.3.6.0.

In II.3.6.A, it is shown that (III.1') implies (III.2').

In II.3.6.B, it is shown that (III.1') is true.

It can be seen that if Theorem 3 is established for the case of $\{g^{tt}, g^o\}$ empty, then Theorem 3 is also true if (1) g^{tt} is not empty, and/or (2) g^o is not empty but ${}_o\phi^I(x^I, y^t)$ has a non-negative saddle-point at $(\bar{x}^I, \bar{y}^t)$, since in either case $\bar{x}$ remains unchanged and the additional terms in the modified Lagrangian expression vanish at $\bar{y}$ (cf. eqs. (B.6) and (C.6'") in II.3.4).

Hence, with no loss of generality, we may henceforth assume $\{g^{tt}, g^o\}$ to be empty, i.e.,

(1) $$g^{\{1\}} = g^*.$$

---

22. The desirability of explicit treatment of the linear case was emphasized by Dr. Masao Fukuoka.

II.3.6.A. (III.1') implies (III.2').

In this section we show that (III.1') implies (III.2'), i.e., that, in a sufficiently small neighborhood, if (III.1') is assumed to be valid and the inequalities $x \geqq 0$, $x \neq \bar{x}$, hold, (III.2) follows. We write $\emptyset$ instead of $_{\eta}\emptyset$ throughout. (III.1'), $x \geqq 0$, $x \neq \bar{x}$, are also assumed throughout II.3.6.A.

Let

$$\xi = x - \bar{x}$$

$$\xi^{i} = x^{i} - \bar{x}^{i}, \qquad i = I, II.$$

(a) First case: $\xi^{II} \neq 0$.

By (I.3") and (N.14**),

(1) $$\bar{\emptyset}_{x} \cdot \xi = \bar{\emptyset}_{x^{I}} \cdot \xi^{I} + \bar{\emptyset}_{x^{II}} \cdot \xi^{II} < 0.$$

But then (III.2) follows from the well known ('Fréchet') property of differentials[23] which, as applied to the present case, states that, given any $\sigma > 0$, there exists an $\epsilon > 0$ such that,[24]

(2) $$\left| \frac{1}{|\xi|} \left(\emptyset(x,\bar{y}) - \emptyset(\bar{x},\bar{y}) - \bar{\emptyset}_{x} \cdot \xi\right) \right| < \sigma$$

if

(3) $$|\xi| < \epsilon .$$

Choose

(4) $$\sigma = - \frac{1}{|\xi|} \bar{\emptyset}_{x}.$$

which is positive by (1).

---

23. See Hille [10, p. 72], Definition 4.3.4, eq. (iii).

24. If a is a real number, $|a|$ denotes its absolute value. If $a = \{a_1, a_2, \ldots, a_K\}$ with the a_k real, $|a| = \left|(\sum_{k=1}^{k} a_k^2)^{1/2}\right|$ denotes the 'length' of a.

Then, for a sufficiently small $|\xi|$, we have (by (2))

(5) $$\left|\frac{1}{|\xi|}(\phi(x,\bar{y}) - \phi(\bar{x},\bar{y})) + \sigma\right| < \sigma$$

which implies

(6) $$\frac{1}{|\xi|}(\phi(x,\bar{y}) - \phi(\bar{x},\bar{y})) < 0$$

and hence (III.2).

If x^I is empty, this completes the proof of the Theorem 3, since $x \neq \bar{x}$ then implies $\xi^{II} \neq 0$. If x^I is not empty, we must consider the

(b) <u>Second case</u>: $\xi^{II} = 0$.

Since it is assumed that $x \neq \bar{x}$, $\xi^{II} = 0$ implies

(7) $$\xi^I \neq 0.$$

In virtue of the existence of the second derivatives of ϕ with regard to x^I (by definition of ϕ, and the assumptions concerning the second derivatives of $\underline{f}$ and $\underline{g}$ with regard to x^I) we have, by Taylor's Theorem,

(8) $$\phi(x,\bar{y}) - \phi(\bar{x},\bar{y}) = \bar{\phi}_{x^I} \cdot \xi^I + \frac{1}{2}(\xi^I)' \tilde{\phi}_{x^I x^I} \xi^I$$

where $\tilde{\phi}_{x^I x^I}$ denotes $\phi_{x^I x^I}$ evaluated at $x = \tilde{x}$, $\tilde{x} = \bar{x} + \theta\xi$, $0 < \theta < 1$. It now suffices to note that $(\xi^I)' \phi_{x^I x^I} \xi^I$ is negative at $\bar{x}$ (since (III.1') is assumed to hold and its hypotheses are satisfied) and continuous in the neighborhood of $\bar{x}$ (by the hypotheses of the theorem concerning the second derivatives of $\underline{f}$ and $\underline{g}$), so that, for a sufficiently small $|\xi^I|$, $(\xi^I)' \tilde{\phi}_{x^I x^I} \xi^I < 0$. Since $\bar{\phi}_{x^I} \cdot \xi^I = 0$ by (N.14**), (III.2) follows.

II.3.6.B.

(III.1) <u>holds if</u> x^I <u>is non-empty</u>.

B.1. <u>First case:</u> g^* <u>empty</u>.

By eq. (1) in II.3.6.0, $g^{[1]}$ is also empty. Hence, by (1) in Theorem 1,

(1') $\bar{y}(\eta) = \bar{y}^{[2]}(0) = 0.$

and, using (N.12),

(1") ${}_{\eta}\phi(x, \bar{y}(\eta)) = f(x).$

Since g^* empty, we have $M^* = 0$, and, therefore, the definition (D.4) of $\underline{q}$ applies, so that (since x^* is empty but x^I is not) t^{**} is not empty and

(2) $\bar{q}_{t^{**}t^{**}} = \bar{f}_{x^{**}x^{**}} = \bar{\phi}_{x^I x^I}$.

(1") and (2), together with the third regularity condition $\overline{3}/$ (eq. (D.19)), yield (III.1) for a sufficiently small neighborhood of $\bar{x}$.

B.2. <u>Second case:</u> g^* <u>non-empty</u>.

Write

B.2.1.

(0) $\psi(t, y) = \phi(r(t), \bar{x}^{II}, y)$

where r(t) is defined in (D.14). (Where it is desired to indicate the dependence of ψ on η, we may write ${}_{\eta}\psi$ instead of ψ.)

Then, by (D.13) (i.e., $\overline{1}/$), we have

(2) $\bar{\psi}_{tt} = \psi_{tt}\big|_{t=\bar{t}} = (\bar{H}^{-1})' \; {}_{\eta}\bar{\phi}_{x^I x^I} \; \bar{H}^{-1}, \; \bar{t} = \{h^*(\bar{x}), \bar{x}^{**}\},$

since ${}_{\eta}\bar{\phi}_{x^I} = {}_{o}\bar{\phi}_{x^I} = 0$ by (I.3") and (N.14**).

We shall now show that (III.1) is implied by

(3) $\tau' \bar{\psi}_{tt} \tau < 0 \quad \text{if } \tau^{(21)} \geqq 0 \quad \text{and} \quad \tau \neq 0$

where the partitioning of τ corresponds to that of $\underline{t}$. (Sections B.2.2 - B.2.6 are then devoted to showing that (3) holds.)

To see that (3) implies (III.1), let x^I satisfy the inequalities $x^{(21)} \geqq 0$, $x^I \neq \bar{x}^I$. Choose

$$(4) \qquad \begin{bmatrix} t^* \\ \\ t^{**} \end{bmatrix} = \tau = \bar{H}(x^I - \bar{x}^I) = \begin{bmatrix} \bar{h}^*_{x^*} & \bar{h}^*_{x^{**}} \\ 0 & I \end{bmatrix} \begin{bmatrix} x^* - \bar{x}^* \\ \\ x^{**} - \bar{x}^{**} \end{bmatrix}$$

Since, by (D.13), $\bar{H}$ is non-singular, $x^I \neq \bar{x}^I$ implies $\tau \neq 0$. Also, (4) yields

$$(5') \qquad \tau^{**} = x^{**} - \bar{x}^{**} ,$$

hence, in particular,

$$(5'') \qquad \tau^{(21)} = x^{(21)} - \bar{x}^{(21)} .$$

But

$$(6) \qquad \bar{x}^{(21)} = 0,$$

since $x^{(21)}$ is a component of $x^{(2)}$ by (N.13), and $\bar{x}^{(2)} = 0$ by (N.6)

Hence

$$(7) \qquad \tau^{(21)} = x^{(21)}$$

and thus $x^{(21)} \geqq 0$ implies $\tau^{(21)} \geqq 0$.

Having shown that the hypotheses of (III.1) imply those of (3), we see that the hypotheses of (III.1), together with the validity of the assertion in (3), yield

$$(8) \qquad \tau' \bar{\Psi}_{tt} \tau < 0$$

But, using in succession (4), (2), and simplifying, we have

(9) $$\tau' \Psi_{tt} \tau = (x^I - \bar{x}^I)' \bar{H}' \bar{\Psi}_{tt} \bar{H}(x^I - \bar{x}^I)$$

$$= (x^I - \bar{x}^I)' \bar{H}' (\bar{H}^{-1})' {}_{\eta}\phi_{x^I x^I} \bar{H}^{-1} \bar{H}(x^I - \bar{x}^I)$$

$$= (x^I - \bar{x}^I)' {}_{\eta}\bar{\phi}_{x^I x^I} (x^I - \bar{x}^I).$$

(8) and (9) yield the conclusion of (III.1). Thus it has been established that (3) implies (III.1). It remains to be shown that (3) is valid. (This is done in sections B.2.2 - B.2.6.)

B.2.2.

It is convenient to write $\bar{\Psi}_{tt}$ in the partitioned form

(1) $$\bar{\Psi}_{tt} = \begin{bmatrix} \bar{\Psi}_{t^*t^*} & \bar{\Psi}_{t^*t^{**}} \\ \bar{\Psi}_{t^{**}t^*} & \bar{\Psi}_{t^{**}t^{**}} \end{bmatrix} = \begin{bmatrix} A & B \\ B' & C \end{bmatrix}$$

where t^{**} (but not t^*) may be empty. (The case of t^* empty was treated in B.1.)

B.2.3.

It will now be shown that A (i.e., $\bar{\Psi}_{t^*t^*}$, cf. eq. (1) in B.2.2), which depends on η, can be made negative definite by a suitable choice of η.

Recalling that $\mathcal{M}^*$ denotes the set of indices of the components of g^*, and using (N.9) and (D.9-1), we see that, for $m \in \mathcal{M}^*$,

(1) $${}_{\eta_m}p_m(x) = 1 - t_m^{1+\eta_m}$$

where t_m is a component of t^*.

Since, by Theorem 1 and eq. (1) in II.3.6.0,

(2) $\quad \bar{y}_m(\eta) = 0 \qquad \text{for } m \in \mathcal{M} \sim \mathcal{M}^*$

we have, from the definitions of Ψ, $\underline{q}$, and $_\eta\emptyset$ (eqs. (1) in B.2.1, (D.17), and (N.12), respectively), and the preceding relations (1) and (2), the equality

(3) $\quad \Psi(t, \bar{y}(\eta)) = q(t) + \sum_{m \in \mathcal{M}^*} [\bar{y}_m(\eta)](1 - t_m^{1+\eta_m})$.

Writing

(4) $\quad F = \bar{q}_{t=t^*}$,

we have, from (3) and the definition of A in B.2.2,

(5) $\quad A = F - D$

where $D = \|d_{m,m'}\|$, $m \in \mathcal{M}^*$, $m' \in \mathcal{M}^*$, is a diagonal matrix (i.e.,

(6') $\quad d_{m,m'} = 0 \qquad \text{for } m \neq m'$)

with

(6") $\quad d_{m,m} = [\bar{y}_m(\eta)]\,(1+\eta_m)\,\eta_m = \bar{y}_m(0)\,\eta_m, \qquad m \in \mathcal{M}^*$,

where the second equality follows from (1.3').

Let λ denote the largest characteristic root of F. Since, by the second regularity condition ($\nabla 2$), $\bar{y}_m(0) > 0$ if $m \in \mathcal{M}^*$, we may choose η_m^o, for each $m \in \mathcal{M}^*$, to be a positive even integer satisfying

(7') $\quad \eta_m^o > \lambda / \bar{y}_m(0)$,

so that

(7") $\quad \min_{m \in \mathcal{M}^*} d_{m,m} > \lambda \qquad \text{for all acceptable } \eta_m \geq \eta_m^o$.

Then, for any $t^* \neq 0$, and each acceptable $\eta_m \geq \eta_m^o$, we have

(8') $$t^{*\prime} F t^* \leq \lambda t^{*\prime} t^* = \sum_{m \in \mathcal{M}^*} t_m^2 < \sum_{m \in \mathcal{M}^*} d_{m,m} t_m^2$$

$$= t^{*\prime} D t^* ,$$

i.e.,

(8") $t^* \neq 0$ implies $t^{*\prime}(F-D)\, t^* < 0$ for all sufficiently large acceptable η,

or

(8'") A is negative definite for all sufficiently large acceptable η.

(8'") suffices to establish B.2.1, eq. (3), and, therefore, (III.1), if t^{**} is empty.

B-2.4. Now assume t^{**} not empty. Write

(1) $$P = \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix}$$

and

(2) $$[illegible]_{\eta} = P'_{\eta} \overline{F}_{tt} P.$$

Then methods used in B.2.1 to show that B.2.1 eq. (3) implies (III.1) can be used to show that

(3) w'_{η} [illegible] $w < 0$ for $w \neq 0$, $w^{(21)} > 0$

implies B.2.1 eq. (3). [For $P^{-1} = \begin{bmatrix} I & A^{-1}B \\ 0 & I \end{bmatrix}$, and, like its analogue $\overline{H}$ in B.2.1, performs an identity transformation on t^{**}, so that the condition $t^{(21)} \geq 0$ is transformed into the condition $w^{(21)} \geq 0$.] It remains to establish (3). Now from (2), (1), and B.2.2. eq. (1), we have

$$\eta^{-}\Omega = \begin{vmatrix} A & 0 \\ 0 & C-B'A^{-1}B \end{vmatrix},$$

so that

$$w'\eta^{-}\Omega w = w^{*\prime} Aw^{*} + w^{**\prime}(C-B'A^{-1}B)w^{**}.$$

Now, by B.2.3 eq. (8'"), we may take A as negative definite, and hence, to establish (3), it will suffice to show that

(4) $\quad = w^{**\prime}(C-B'A^{-1}B)w^{**} < 0 \quad$ if $w^{**} \neq 0$, $w^{(21)} \geq 0$.

This is done in B.2.6 after an auxiliary result is obtained in B.2.5.

B.2.5. It will now be shown that the norm[25] of A^{-1} can be made arbitrarily small by choosing η sufficiently large.

D^{-1} is a diagonal matrix whose non-zero elements approach zero for η large; hence $N(D^{-1})$ can be made arbitrarily small for η large. Then $N(D^{-1}F) \leq N(D^{-1})\,N(F)$ approaches 0. For η sufficiently large, $N(D^{-1}F) < 1$; it follows that[26]

$$N\left[(D^{-1}F-I)^{-1}\right] < \frac{1}{1-N(D^{-1}F)}.$$

Since

$$A = F - D = D(D^{-1}F-I),$$

---

25. It does not matter which of the many norms is used; cf. Bowker [6]. Note that, denoting by $N(X)$ the norm of the matrix X, we have $N(A+B) \leq N(A) + N(B)$, $N(AB) \leq N(A)\,N(B)$; if all the elements of a matrix approach 0, so does its norm. If I denotes the identity matrix, $N(I) = 1$.

26. Waugh [7, p. 148]. Let B be a square matrix such that I-B is non-singular. In view of the identity

$$(I-B)^{-1} = I + (I-B)^{-1}B$$

the properties of the norm yield

$$N\left[(I-B)^{-1}\right] \leq 1 + N\left[(I-B)^{-1}\right]N(B)$$

which, for $N(B) < 1$, yields

$$N\left[(I-B)^{-1}\right] \leq \frac{1}{1-N(B)}.$$

it follows that $$A^{-1} = (D^{-1}F-I)^{-1}D^{-1},$$

and hence $$N(A^{-1}) < N(D^{-1})\, N\left[(D^{-1}F-I)^{-1}\right] \leq \frac{N(D^{-1})}{1-N(D^{-1}F)}$$

which can be made arbitrarily small for η large.

B.2.6. Consider now the quadratic form [illegible] in B.2.4 eq. (4). By B.2.2 and B.2.3 eq. (3),

$$C = \overline{\psi}_{t^{**}t^{**}} - \overline{q}_{t^{**}t^{**}}.$$

Hence

The third regularity condition ([3], (D.19")) implies

(1) $w^{**\prime}\, C\, w^{**} < 0$ if $w^{**} \neq 0,\ w^{(21)} \geq 0.$

As shown in B.2.5,

$$N(B'A^{-1}B) \leq N(B')\, N(A^{-1})\, N(B) = N(A^{-1}) \left[N(B)\right]^2$$

can be made arbitrarily small by choosing a large enough η. Now

(2) $\left| w^{**\prime}B'A^{-1}Bw^{**} \right| \leq N(B'A^{-1}B)\, w^{**\prime}w^{**}$,

since the characteristic roots of a matrix are bounded in absolute value by its norm.

Also, denoting by μ the maximum of $w^{**\prime}Cw^{**}$ subject to $w^{**\prime}w^{**} = 1$, $w^{(21)} \geq 0$,

we have

(3) $w^{**\prime}Cw^{**} \leq \mu\, w^{**\prime}w^{**}$

and, by (1),

(4) $\mu < 0.$

With the aid of (2),

(5) [illegible] $< \left[\mu + N(B'A^{-1}B)\right] w^{**\prime}w^{**}$ if $w^{**} \neq 0,\ w^{(21)} \geq 0.$

By choosing η sufficiently large, so that

(6) $\mu + N(B'A^{-1}B) < 0,$

we establish B.2.4 eq. (4) which, in turn, yields B.2.4 eq. (3), B.2.1 eq. (3), (III.1), and hence Theorem 3.

III. Game-Theoretical and Economic Interpretation of the Modified Lagrangian Approach.

III.1. Reasons for Using the Modified Lagrangian Approach

The obvious incentive for using the modified Lagrangian approach (i.e., using ${}_{\eta}\phi$ with a suitably large η instead of $\phi = {}_{0}\phi$) arises in cases where the convexity hypotheses of Theorem 3 i. [1] are not satisfied. However, even when ${}_{0}\phi$ does have a saddle-point, it may be preferable to use the modified Lagrangian approach (η with some positive components). Such is the case when certain gradient procedures for reaching the saddle-point are to be used (cf. [3]) and it is essential that the matrix ${}_{\eta}\phi_{x^I x^I}$ be negative definite. The latter condition would not, for instance, be satisfied by the unmodified Lagrangian expression (${}_{0}\phi$) in the linear case (i.e., where both $f(x)$ and $g(x)$ are linear in $\underline{x}$), although the convexity conditions of Theorem 3 in [1] are satisfied and a saddle-point does exist.[27]

III.2. Game Interpretation

Whenever ${}_{\eta}\phi$ possesses a saddle-point, a game analogous to g_{ϕ} in I.3 may be set up, with ${}_{\eta}\phi$ replacing ϕ. This game may be denoted by $g({}_{\eta}\phi)$. Bearing in mind the local nature of the theorems in II, $g({}_{\eta}\phi)$ still has the important property $\bar{x} = \hat{x}$.

Similarly, an analogue of g^*_{ϕ} may be constructed by using the appropriate terms of ${}_{\eta}\phi$ instead of those of ϕ. The economic interpretation of such a game, as well as the analogue of $g^{*D}_{\phi A}$ (in I.6) will be considered in III.3.

---

27. The gradient methods of the type described in [3] result in constant amplitude oscillations (rather than convergence) when applied to the unmodified Lagrangian expression for the linear case. Cf. Samuelson, [8], pp. 17-22. The desire to remedy this motivated the present work to a considerable extent.

III.3 Economic Interpretation; Decentralization

Consider first the analogue of the allocation game $g^*_{\emptyset A}$ (cf. I.5) in the modified Lagrangian approach. If the expressions derived from $_\eta\emptyset$ are substituted for those derived from $\emptyset$, the nature of the game is altered in only one essential respect: the terms on which the custodian sells primary commodities to the manager. It is easily seen that the selling price (per unit) of the m-th primary commodity is now given by $y_m\left[h_m(x)\right]^{\eta_m}$.

The dependence of the price on the amount purchased is a familiar phenomenon in the economics of "imperfect" competition. There it typically arises under conditions of "increasing returns" corresponding to those where the convexity conditions of Theorem 3 in [1] fail to hold; it also arises under conditions of "constant returns" corresponding to the linear homogeneous case (cf. eq.(12) in I.6) where a saddle-point exists but $\partial\bar{\emptyset}_{x^I x^I}$ is not negative-definite. These facts suggested initially the possibility of the modified Lagrangian approach.

There are a number of analogues of equations I.6 (14-15). One example is the following:

(1) $$_\eta\pi^{y_m}(x,y) = -\,y_m h_m + p_m h_m(x), \qquad m = 1,\ldots,M,$$

(2) $$_\eta\pi^{x_n}(x,y) = \sum_{p=1}^{P} a_p\, k_p^n(x_n) - \sum_{m=1}^{M} p_m\, h_m^n(x_n), \qquad n = 1,\ldots,N,$$

where

(3) $$p_m = y_m\left[h_m(x)\right]^{\eta_m}.$$

Note that p_m depends on, among other things, x_n, so that the $n^{\underline{th}}$ manager must take into account the impact of his decisions on the price of the $m^{\underline{th}}$ primary commodity in this decentralized economy.

The decentralization of the system consists in the fact that a custodian need only know his own ($m^{\underline{th}}$ component) of h and the net demand for that commodity (which helps determine p_m), while the $n^{\underline{th}}$ manager need know only the functions k_p^n ($p = 1,\ldots,p$), h_m^n($m = 1,\ldots, M$), and the net demand for his product $h_m(x)$.

Appendix 1.[28]

Let the first regularity condition ($\nabla 1$) hold. Consider $\bar{x}$ such that,[29]

$$(1)\qquad g^{[1]}(\bar{x}) = 0 \quad , \quad g^{[2]}(\bar{x}) > 0$$
$$\bar{x} \geqq 0$$

and x such that

$$(2)\qquad \bar{g}_x^{[1]}(x-\bar{x}) \geqq 0 \ ,$$
$$x^{(2)} - \bar{x}^{(2)} \geqq 0 \ .$$

Define now the function $g^{\#}$ of $\bar{x}$ by

$$(3)\qquad g^{\#}(x) = \left\{ g^{t}(x),\ x^{**},\ x^{II} \right\}$$

where x^{**} is defined in II.3.4.D.

[Assuming g^{o} to be empty (cf. II.3.4.C), $g^{\#}$, like $\underline{x}$, has N dimensions.]

It follows that

$$(4)\qquad g_x^{\#} = \begin{bmatrix} g^{t}_{x^{*}} & g^{t}_{x^{**}} & g^{t}_{x^{II}} \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}$$

and hence

$$(5)\qquad \left| \bar{g}_x^{\#} \right| = \left| \bar{g}_{x^{*}}^{\#} \right| \neq 0 \ .$$

Consider now the relation which associates with a real number α the values $\dot{x}$ of $\underline{x}$ for which the equation

$$(6)\qquad g^{\#}(\dot{x}) = g^{\#}(\bar{x}) + \alpha \bar{g}_x^{\#} (x-\bar{x})$$

---

28. This appendix parallels Lemma 76.1, in Bliss [11].

29. In this appendix all inequalities are to be interpreted in the sense of (N.1.4).

is satisfied. In virtue of the implicit function theorem, for sufficiently small values of α (6) defines $\bar{\bar{x}}$ as a (single-valued) differentiable function of α, say

(7) $\bar{\bar{x}} = \Psi_1(\alpha)$,

such that

(8) $\Psi_1(0) = \bar{x}$.

Differentiating (6) with respect to α and setting $\alpha = 0$, we have

(9) $\bar{g}_x^{\#} \Psi_1'(0) = \bar{g}_x^{\#} (x-\bar{x})$

and hence, because of (5),

(10) $\Psi_1'(0) = x-\bar{x}$.

We shall now show that

(11) $\Psi_1(\alpha) \in C_g$ for $\alpha \geq 0$, α sufficiently small.

By (6), (1), (2)

(12) $g^t(\bar{\bar{x}}) = \alpha\, \bar{g}_x^t (x-\bar{x}) \geq 0$ for $\alpha \geq 0$.

It follows that

(13) $g^{[1]}(\bar{\bar{x}}) \geq 0$ for $\alpha \geq 0$.

which together with

(14) $g^{[2]}\left[\Psi_1(\alpha)\right] \geq 0$ for α sufficiently small,

yields

(15) $g\left[\Psi_1(\alpha)\right] \geq 0$ for $\alpha \geq 0$ sufficiently small.

Now, since x* is a subvector of $x^{(1)}$, $x^{(2)}$ is a subvector of $\{x^{**}, x^{II}\}$, hence (2) and (6) imply

(16) $\bar{x}^{(2)} = \Psi_1^{(2)}(\alpha) = \bar{x}^{(2)} + \alpha\,(x^{(2)} - \bar{x}^{(2)}) \geq 0 \quad$ for $\alpha \geq 0$

which, together with

(17) $\bar{x}^{(1)} = \Psi_1^{(1)}(\alpha) \geq 0 \quad$ for a sufficiently small

yields

(18) $\Psi_1(\alpha) \geq 0 \quad$ for $\alpha \geq 0$, α sufficiently small.

In turn, (15) and (18) yield (11).

Now let us interpret "α sufficiently small" as $0 \leq \alpha \leq \lambda$ where $\lambda > 0$ and define the function Ψ by

(19) $\Psi(\theta) = \Psi_1(\lambda\,\theta) \quad$ for all $0 \leq \theta \leq 1$.

Then

$$\Psi(0) = \bar{x}$$

(20) $\Psi'(0) = \lambda\,\Psi_1'(0) = \lambda\,(x - \bar{x}) \qquad (\lambda > 0)$

$\Psi(\theta) \in C_g \quad$ for $\quad 0 \leq \theta \leq 1$

Since (20) are precisely the requirements of C.Q., it has been shown that $\bigtriangledown\!1$ implies C.Q.

Appendix 2.

We shall now show that, if the first two regularity conditions hold and in a neighborhood of $\bar{x}$, $\underline{f}$ and $\underline{g}$ are assumed to possess continuous derivatives of second order with regard to x^I, (D.18) is valid.

Let x^{**} be non-empty. Then, writing

$$(1)\quad \begin{cases} -1) & \bar{t}^{*} = h^{*}(\bar{x}) = 1 \quad \text{(a vector 1's)}, \\ -2) & \bar{t}^{**} = \bar{x}^{**}, \end{cases}$$

we have, using Taylor's theorem,

$$(2)\qquad q(\bar{t}^{*}, t^{**}) - q(\bar{t}^{*}, \bar{t}^{**}) = \bar{q}_{t^{**}} \cdot (t^{**} - \bar{t}^{**})$$

$$+ \frac{1}{2}(t^{**} - \bar{t}^{**})' \tilde{q}_{t^{**}t^{**}} (t^{**} - \bar{t}^{**})$$

where '‾' over a symbol denotes the evaluation at $t = \bar{t}$, while '~' over a symbol denotes evaluation at $t = \tilde{t}$, $\tilde{t} = \bar{t} + \theta(t^{**} - \bar{t}^{**})$, $0 < \theta < 1$. Now suppose it has been shown that (a) $q(\bar{t}^{*}, t^{**})$, has, as a function of t^{**}, subject to the constraint $t^{(21)} \geq 0$, a local maximum at $t^{**} = \bar{t}^{**}$, and (b) $\bar{q}_{t^{**}} = 0$. From (a) it follows that, in a sufficiently small neighborhood, the left member of (2) is non-positive if $t^{(21)} \geq 0$. But then, using (b), we see that the quadratic form in the right member of (2) is non-positive. Since, by hypothesis, $q_{t^{**}t^{**}}$ is a continuous function of t^{**}, we have, for $t^{(21)} \geq 0$, and in a sufficiently small neighborhood of $\bar{t}$,

$$(3)\qquad (t^{**} - \bar{t}^{**})\, \bar{q}_{t^{**}t^{**}} (t^{**} - \bar{t}^{**}) \leq 0$$

which is the desired result (D.18). Hence it remains to prove (a) and (b).

(a) <u>$q(\bar{t}^*, t^{**})$, has as a function of t^{**}, subject to $t^{(21)} \geq 0$, a local maximum at</u> $t^{**} = \bar{t}^{**}$.

It follows from the remarks at the beginning of II.3.4.C that $f(x^I, 0)$, as a function of x^I, has a local maximum at $x^I = \bar{x}^I$, subject to the constraints

$$g^t(x^I, 0) = 0$$

(4)

$$x^{(21)} \geq 0 \; .$$

Hence, the subject to the same constraints, $q(t)$ has a local maximum at $\bar{t}$. Now we must distinguish the two ways in which the "milder" (C.8') second regularity condition ($\nabla 2$) may be satisfied.

(<u>1</u>) $_0\phi^I\,(x^I, y^t)$ has a non-negative saddle-point at $(\bar{x}^I, \bar{y}^t)$. I.e., locally, since $\bar{y}^\circ = 0$ by (C.6'"),

(5) $$f(x^I,0) + \bar{y}^* \cdot g^*(x^I,0) \leq f(\bar{x}^I,0) + \bar{y}^* \cdot g^*(\bar{x}^I,0))$$

for all x^I such that $x^{(21)} \geq 0$.

But $g^*(\bar{x}^I, 0) = 0$ because of (4), and $g^*(x^I,0)$ in the left member of (5) vanishes for $t^* = \bar{t}^*$. Hence (5) yields, locally and for $t^{(21)} \geq 0$,

(6) $$f(r^*(\bar{t}^*, t^{**}), t^{**}, 0) \leq f(r^*(\bar{t}^*, \bar{t}^{**}), \bar{t}^{**}, 0)$$

which means precisely that $q(\bar{t}^*, t^{**})$ has a local maximum at $\bar{t}^{**}$ subject only to $t^{(21)} \geq 0$.

(<u>2</u>) g° is empty.

In this case (4) is equivalent to

-1) $$g^*(x^I, 0) = 0$$

(7)

-2) $$x^{(21)} \geq 0.$$

But (7-1) is necessarily satisfied if $t^* = \bar{t}^*$ and hence can be disregarded. Since $q(t)$ was seen to have a local maximum at $\bar{t}$ subject to (4), it follows that $q(\bar{t}^*, t^{**})$ will have a local maximum at $\bar{t}^{**}$ subject only to $t^{(21)} \geq 0$.

(b) $\bar{q}_{t**} = 0.$

We have

(8) $\bar{q}_{t**} = \bar{f}_{x*} \bar{r}^*_{t**} + \bar{f}_{x**} .$

We now evaluate the three expressions on the right-hand side of (8). We start with $\bar{r}^*_{t**}$. Noting that

(9) $g^*\left\{\left[r^*(\bar{t}^*, t^{**}), t^{**}\right], 0\right\} = 0$ for all t^{**},

we obtain by differentiation with respect to t^{**}, using (D.9-1) and (D.16), and evaluating at $t = \bar{t}$,

(10) $\bar{g}^*_{x*} \bar{r}^*_{t**} + \bar{g}^*_{t**} = 0$

in virtue of (C.10) ($\nabla 1$) this can be solved yielding

(11) $\bar{r}^*_{t**} = -(\bar{g}^*_{x*})^{-1} \bar{g}^*_{x**} .$

To find $\bar{f}_{x*}$, $\bar{f}_{x**}$, we write the condition that ${}_0\phi^I_{x^I} = 0$ (eq. (C.4-1)) in the form

(12) $$\begin{cases} \bar{f}_{x*} + \bar{g}^*_{x*} \bar{y}^* = 0 \\ \bar{f}_{x**} + \bar{g}^*_{x**} \bar{y}^* = 0. \end{cases}$$

(The terms involving g^0 vanish, of course.)

Substituting (12) and (11) into (8), we have

(13) $$\begin{aligned} \bar{q}_{t**} &= (-\bar{y}^* \cdot \bar{g}^*_{x**}) + (-\bar{y}^* \cdot \bar{g}^*_{x*}) \left[-(\bar{g}^*_{x*})^{-1} \bar{g}^*_{x**}\right] \\ &= (-\bar{y}^* \cdot \bar{g}^*_{x**}) + (\bar{y}^* \cdot \bar{g}^*_{x**}) \\ &= 0 . \end{aligned}$$

This completes the proof of (D.18).

REFERENCES

[1] H. W. KUHN and A. W. TUCKER, "Nonlinear Programming," J. Neyman (ed.), Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, Berkeley and Los Angeles, University of California Press (1952), pp. 481-492.

[2] T. C. KOOPMANS, "Analysis of Production as an Efficient Combination of Activities," T. C. Koopmans (ed.), Activity Analysis of Production and Allocation, Cowles Commission Monograph No. 13, New York, John Wiley & Sons, 1951, pp. 33-97.

[3] K. J. ARROW and L. HURWICZ, "A Gradient Method for Approximating Saddle-Points and Constrained Maxima," P-223, 13 June 1951, The RAND Corporation (hectographed).

[4] G. B. DANTZIG, "Maximization of a Linear Function of Variables Subject to Linear Inequalities," in T. C. Koopmans (ed.), op. cit., pp. 339-347.

[5] P. A. SAMUELSON, Foundations of Economic Analysis, Cambridge, Mass., Harvard University Press, 1948.

[6] A. H. BOWKER, "On the Norm of a Matrix," Annals of Mathematical Statistics, Vol. 18 (1947), pp. 285-288.

[7] F. V. WAUGH, "Inversion of the Leontief Matrix by Power Series," Econometrica, Vol. 18 (1950), pp. 142-154.

[8] P. A. SAMUELSON, "Market Mechanisms and Maximization," P-69, 28 March 1949, The RAND Corporation (hectographed).

[9] L. HURWICZ, "Programming in General Spaces," Cowles Commission Discussion Paper: Economics No. 2109, July 29, 1954 (hectographed).

[10] E. HILLE, Functional Analysis and Semi-Groups, New York, American Mathematical Society Colloquium Publications, Vol. XXXI, 1948.

[11] G. A. BLISS, Lectures on the Calculus of Variations, Chicago, The University of Chicago Press, 1946.